DOCUMENT RESUME

ED 185 931 HE 012 639

AUTHOR Stimac, Michele
TITLE Humanism Begins at Home.
PUB DATE .6
NOTE 8p.
JOURNAL CIT Pepperdine Commentator: v1 n1 p7-12 Fall 1976

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *College Faculty; College Students; *Evaluation Criteria; Higher Education; Reliability; Research Problems; *Student Evaluation of Teacher Performance; *Teacher Effectiveness; Teacher Improvement; *Teacher Morale; Validity

ABSTRACT

The trend of student evaluation of college faculty performance is documented, and implications for humanization of the university are considered. Research in the area of teacher evaluation is cited, and it is proposed that reviews of the literature on student evaluations indicate by and large that student ratings are reliable and valid, even though there is a large body of dissenting opinion. Students have obtained significant influence in evaluation of teaching effectiveness, and administrators are thought to be viewing student evaluation of teacher effectiveness as more reliable and valid than colleague evaluation or informal observation. There still remains the question of how feedback on teacher performance can be used constructively for teacher growth and improvement. Evidence seems to support the notion that teacher morale has declined decidedly on many campuses. It is suggested that if inept methods of evaluation and accountability pressures cause poor morale of faculty members, there is a need for reanalysis of the issues. Clarification of whether evaluation of instructors is to improve teaching effectiveness or to determine who will be promoted or tenured is needed. Student ratings need to be carefully weighed in view of the complexity of the nature of the process. Until further research is accomplished to develop sophisticated and tested evaluation forms, current student evaluation should be kept in perspective. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED185931

HUMANISM BEGINS AT HOME

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Michele Stimac

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

MICHELE STIMAC

Ages of wisdom have dictated a rule that says "charity begins at home." It might not be too far afield to suggest that in an age of humanistic endeavor "humanism begins at home." I speak specifically with respect to the vast and complicated issue of evaluation and accountability demanded of professors in our academic institutions today. I use *humanism* in the sense that "third force" humanistic psychologists might use it. Humanism nurtures special regard and concern for the individual and for what is human even when increasing size and technology dehumanize and depersonalize the pulse of institutional and interpersonal enterprise.

Columbia, 1968, and subsequent student revolts were snowballs that grew to become potential and in some instances real avalanches. As a result, what students think and feel about their educational experiences have become increasingly significant. The campus violence and riots of the late 60's brought institutions up short, made them heed the shouts of students, made them respond to their demands and in some instances even cower at their threats. It did not take long for "student power" to make itself felt.

One area in which the student voice was heard and continues to be heard is instructional evaluation. There are few institutions today that do not have

HE 012 639

some form of student rating of teacher performance. And, though people like Kerlinger contend that "evaluations of professors and their teaching initiated and conducted by students are not an integral part of the instructional process,"[1] student ratings continue to be recognized as reliable and valid criteria for assessment of instruction.

Research in the area of teacher evaluation has been getting increased attention in the last few years. People like John Centra at ETS and Peter Frey at Northwestern University, to mention two, conducted significant attempts at sophisticated research in the area of student ratings of instruction.

Centra began an extensive project in the winter of 1970 to develop a questionnaire to gather student ratings of courses and instruction.[2] The final version of the instrument, which he calls the SIR (Student Instructional Report), was ready for general use by 1972-73.

In its most refined form, the SIR contains items that attempt to rate instructors in the areas that previous research[3] has indicated were most essential to students' learning and achievement gains:

> . . .instructors (who) gave clear explanations, were organized in their lessons, stimulated students' intellectual curiosity, gave interesting presentations of course materials, were attentive to students' reactions, were friendly, and were flexible.[4]

The SIR is currently available for general use through ETS.

Michele Stimac is an Associate Professor and Doctoral Program Counselor in the School of Education. She has taught and counseled at all levels of education and has a special interest in career development. She received her doctorate at Boston University in 1973 and places her focus on humanistic psychology and educaton.

A second study undertaken by Centra had as its primary purpose to examine at length the effects of student feedback on teaching at the college level.[5] Researchers such as Tuckman and Oliver,[6] Gage, Runkel, and Chatterjee,[7] and Bryan[8] had discovered at the high school and elementary level that there were gains in teacher ratings as the result of student feedback. Miller, however, in a study which admittedly had limitations, discovered no significant change among college teachers as a result of feedback.[9] Centra's investigation of student ratings, therefore, focused on college level teaching. He included two groups in his study. For the first group, student ratings only were shown to instructors with little or no interpretation to accompany the ratings, the typical procedure at most colleges. Interpretation of results was made available to the second group.

Centra's study led to the conclusion that student ratings can effect changes in teaching, especially for those teachers who have "unrealistically high (compared to their students' views) opinions of their instructional practices."[10] Festinger's notion of cognitive dissonance may be a factor here.[11]

Centra also found that instructional changes would take place especially if teachers "are provided with information to help them interpret their own results."[12] The study underscores the need for comparative or "normative" data for the teacher.

Frey's research is concerned chiefly with the kinds of information that should be solicited about instruction and how it should be solicited.[13] He believes, too, ". . .that there is generally a positive relationship between student ratings and good teaching but that the strength of this relationship depends critically on the technical sophistication of the rating questionnaire."[14]

Extensive reviews of the literature on student evaluations[15] indicate by and large that student ratings are reliable and valid, even though there is a large body of opinion that believes the opposite.

> Typically, they (the faculty) claim that student ratings are unreliable, that the ratings will favor an entertainer over the instructor who gets his material across effectively, that ratings are highly correlated with expected grades (a hard grader would thus get poor ratings), and that students are not competent judges of instruction since long-term benefits of a course may not be clear at the time it is rated.[16]

Costin and his associates conclude, however, that "students can rate classroom instruction with a

reasonable degree of reliability."[17] They refer to a host of studies such as the research done by Spencer and Aleamoni,[18] Costin,[19] and Guthrie,[20] which shows moderate to high correlations between student rankings of teachers from one year to the next.

Costin and his associates approached the question of validity by citing 20 studies which show either no relationship or a weak relationship between students' grades and their ratings of instructors and courses. This finding refutes one of the chief arguments against the validity of student ratings, that teachers who give high grades receive higher student ratings.

Extensive reviews of the literature on student evaluations[15] indicate by and large that student ratings are reliable and valid, even though there is a large body of opinion that believes the opposite.

Citing French,[21] Crawford and Bradshaw,[22] and Musella and Rusch,[23] Costin and associates point out that the criteria students use in evaluating instruction are consistent and are not unlike the elements most contributive to learning cited earlier by Centra in his SIR instrument:

> Among the most frequently mentioned characteristics were expert knowledge of subject matter, systematic organization of course content, ability to explain clearly, enthusiastic attitude toward the subject, and ability to encourage thought.[24]

They also point to the fact that while a study by Royce suggests that high ratings by students were largely the function of "superficial popularity,"[25] Weaver refutes this contention by reporting that students' ratings of instructors were *not* a product of the popularity "halo."[26]

Centra's review of research answered the typical questions raised by opponents to student evaluation.

> Do students with higher gradepoint average rate instructors differently than students with a lower gradepoint average? Guthrie reports they do not.[27]
>
> Are student ratings reliable? Remmers[28] and Guthrie[29] report that when 25 or more ratings were averaged, means were as reliable as the better mental tests available. Centra, using an intraclass measure of reliability, reports that ratings at Michigan State were similarly consistent for most items.[30]
>
> Do male and female students differ in their ratings instructors? Remmers and Elliott[31] and Spencer[32] indicate they do not.
>
> Is there a relationship between the ratings a student gives the instructor and the grade he gets or expects in the course? Some studies indicate that there is a small but significant relationship between these factors.[33] Other studies report no relationship.[34]
>
> Does the curriculum or the students' year in college have any bearing on their course ratings? Spencer reports they have little or no relationship.[35]
>
> Do ratings of alumni differ from the oncampus ratings of instructors? Drucker and Remmers found that alumni, ten years away, agree (r's between .40 and .68) with students in their ratings of the same instructors.[36]
>
> Do teacher personality measures and student ratings have a relationship? Borg[37] and Bendig[38] found the relationship insignificant. Isaacson, McKeachie and Milholland, however, report that psychology instructors rated high in "general cultural attainment" by their peers tended to receive good ratings from students.[39] Cosgrove supports this latter finding,[40] while Filey, Ryan and Lifshitz indicate that in subjects other than psychology, students may rate other faculty qualities as important.[41]

Thus, research studies of varying levels of sophistication appears to strengthen the case for the value and use of student ratings. Like it or not, students have gained significant influence in evaluation of teaching effectiveness, and administrators are turning to student opinion of teacher effectiveness as more reliable and valid than colleague evaluation or informal observation.

A recent survey of private college academic deans by Seldin and Wakin suggests an increase in the use of student ratings to judge teaching effectiveness.[42] "In identifying major factors in evaluating *overall* faculty performance, practically all deans cited classroom teaching as pivotal." They indicate further that there "is a decline in attention paid to research, publication, public service, and activity in professional societies — the

traditional benchmarks of academic success."[43]

For the traditionalist, this trend has a threatening echo, especially if a balanced perspective is not established and/or maintained. Kerlinger decries student initiated and conducted evaluations because they "are unsystematic, lacking the objectivity and control necessary to adequate assessment." He points to several dangerous consequences of such a process. Instructor hostility, resentment, and distrust, he says, will be aroused because the process is a "threat to the autonomy of the professor" and an invasion of teaching responsibility. Furthermore, those who contend "that student evaluation can heighten the professor's sense of responsibility, increase his motivation, and thus improve instruction" are in error because their argument is based on the notion that external sources of motivation are legitimate and effective.[44]

It has often been contended that a student's motivation to achieve and improve, though perhaps urged and inspired by a significant other or some external agent, ultimately comes from within the student him/herself. Perhaps the same principle applies to the teacher's motivation to improve or desire for effectiveness. It cannot be coerced; rather it must come from within the teacher him/herself.

Does this mean that evaluation, feedback, and the call to accountability have no meaning? Decidedly, not. From its inception with Kurt Lewin, the concept of "feedback" has received increasing respect and its effects have been tested and measured in the whole arena of human relations and behavior change. Student feedback is valuable as a provocative and constructivle means for teacher growth and improvement. The question is how this feedback is to be attained and what will be done with it.

Harnett, in speaking of accountability in higher education, has said, "In many ways, educational accountability and evaluation are essentially the same. Accountability, like evaluation, is aimed at learning about the effect of educational institutions."[45] However, he goes on to point out that "accountabilit, has brought with it the notion of external judgment," and "at least from the reactions of many elementary and secondary school teachers, there is a clear indication that 'accountability' is regarded as a vindictive rather than an affirmative process."[46]

McGahan raises a similar issue when he looks upon accountability as a negative reinforcer. "If we say that someone is accountable we usually mean that 'he must suffer the consequences of his actions.' We hardly ever mean the more positive 'he will profit from the consequences of his action'."[47]

There is enough evidence to support the notion that a decided decline in teacher morale is a current phenomenon on many campuses.

It might be well here to return to the opening statement in this paper; namely, that in an age of humanistic endeavor, humanism begins at home. If evaluation and accountability have failed to produce an atmosphere of growth, health, and productivity, if they have become negative reinforcers and have proven to "alienate professors from a vital part of their work,"[48] then something is awry and our humanistic posture is a facade.

There is enough evidence to support the notion that a decided decline in teacher morale is a current phenomenon on many campuses. Tension, factions, hostility, and mistrust are often prevalent. It behooves us to examine why these exist. If one of the causes of decayed morale is inept methods of evaluation and accountability pressures with all their consequent implications, a quick response is in order. Institutional excellence cannot be maintained in a climate of dying morale.

Therefore, since we are in a time of increased evaluation and accountabililty, there are several questions that need to be continually and carefully examined. For one thing, we must clarify why we are evaluating our instructors. Is it to improve the effectiveness of their teaching or is it to determine who will be promoted, who will get tenure, who will get salary increments? It would seem that our current situation in many institutions lacks clarity of motives and perhaps this lack of clarity muddies the entire process of evaluation. Even the most sophisticated instruments would fail to achieve the right end in this instance.

As Kerlinger indicates, "the basic purpose of any evaluation procedure is to evaluate instruction and not instructors, to determine whether the instructional objectives have been reached and how well they have been reached."[49] Of course, he further maintains that "it makes little difference whether students like or do not like a course or an instructor. It really does not matter if an instructor is interesting or dull. Attaining the course's intellectual and value goals is what matters."[50]

Frey, in his concern for the kinds of information we solicit from students and how we obtain it, makes

clear that the student should be viewed as "an information source rather than an evaluator." One should not tolerate questionnaires containing rating items "worded in terms of observables and statements of opinion or attitude" or that fail to treat "the teaching situation as one having many dimensions" that should be rated separately.[51]

...it would be well to remind ourselves that we are still a distance from adequate evaluative processes and procedures and that we must regard the extant ones in this light.

Finally, it might be well to underscore Frey's poignant contention that

> it is important to take into account the fact that students' perceptions are a product of their own personalities as well as of the teacher's behavior. Thus the impression that a teacher creates depends not only on his own behavior but also on the behavior and expectations of his audience. Any analysis which assumes that teacher ratings depend entirely on the target and are independent of their source will be woefully inadequate.[52]

A great deal of research needs to be done in the area suggested by this last point. If we agree with the perceptual psychologists and phenomenologists, we give credence to the notion that each individual possesses a perceptual field that influences the meanings and conclusions he arrives at. Combs cites an illustrative example:

> Several years ago a friend of mine was driving in a car at dusk along a Western road. A globular mass, about two feet in diameter, suddenly appeared directly in the path of the car. A passenger screamed and grasped the wheel attempting to steer the car around the object. The driver, however, tightened his grip on the wheel and drove directly into the object. The behavior of both the driver and the passenger was determined by his own (perceptions). The passenger, an Easterner, saw the object in the highway as a boulder and fought desperately to steer the car around it. The driver, a native Westerner, saw it as a tumbleweed and devoted his efforts to keeping his passenger from overturning the car.[53]

Again, all of this is not to suggest that student ratings have no reliability or validity and that they should be dismissed. It is only to suggest that they need to be carefully weighed in light of the complexity of the entire nature of the process and that, for the sake of humanism, we do not trample on the evaluatee by failing to recognize this complexity. Although research mounts and studies multiply, it would be well to remind ourselves that we are still a distance from adequate evaluative processes and procedures and that we must regard the extant ones in this light.

If we do not believe this, we need only look at data such as the kind Seldin and Wakin report about deans' appraisals of evaluations: ". . .they feel evaluation on their campus is too subjective, and that they have not succeeded in working out satisfactory procedures and methods."[54] While some institutions have achieved a degree of sophistication in their evaluaton process, there are still some whose deans admit, "I am sorry to report that we have not made progress (in teaching evaluation)"; "In this my nineteenth year as a faculty member and my third as academic dean, I am still unable to suggest a truly valid way of evaluating faculty members"; and "We are not satisfied with this (evaluation form). We are trying to improve it from time to time. . .but nothing has been done toward final revision."[55]

To institutions whose deans could make similar admissions, caution and humanistic admonition are in order. Destroy crude and illegitimate rating forms. Investigate the educational field and unearth sophisticated tried-and-tested forms. Until further research is accomplished, keep current student evaluation in perspective. Preserve and nurture the "human professor" while the mechanistic instruments are refined. Finally, and perhaps most importantly, learn from the field of psychometry that an instrument is more often than not only as good as its interpreter.

Conscientious educators would not indulge in administration of an instrument designed to measure student performance without following up its administration with a sophisticated interpretation to the student in a counseling session. Therefore, we must ask who interprets for the teacher the meaning of student ratings? Is there an ongoing counseling process for teachers and professors reflecting a warm, empathic climate and an atmosphere of positive reinforcement? Or must this kind of question be asked only with tongue in cheek?

Humanism begins at home. What we prescribe for our students, we might do well to prescribe for ourselves.

REFERENCES

1. E. Kerlinger, "Student Evaluation of University Professors," *School and Society*, October 1971, p. 353.

2. J. A. Centra, "The Student Instructional Report: Its Development and Uses," SIR Report Number 1, Educational Testing Service, Princeton, N.J., 1972.

3. D. Soloman, W. E. Bezdek, and L. Rosenberg, *Teaching Styles and Learning*, Center for the Study of Liberal Education for Adults, Chicago, 1963; W. R. Mann, "Changes in the Level of Attitude Sophistication of College Students as a Measure of Teacher Effectiveness," *Dissertation Abstracts*, vol. 29, no. 6A, pp. 2443-2444, 1969; and S. A. Cohen and W. G. Berger, "Dimensions of Students' Ratings of College Instructors Underlying Subsequent Achievement on Course Examinations," *Proceedings of the 78th Annual Convention of the American Psychological Association*, Washington, D. C., 1970, *APA*, vol. 6, pp. 605-606, 1971.

4. Centra, op. cit., p. 10.

5. J. A. Centra, "Two Studies on the Utility of Student Ratings for Improving Teaching," SIR Report Number 2, Educational Testing Service, Princeton, N. J., 1972.

6. B. W. Tuckman and W. F. Oliver, "Effectiveness of Feedback to Teachers as a Function of Source," *Journal of Educational Psychology*, vol. 59, pp. 297-301, 1968.

7. N. L. Gage, P. J. Runkel, and B. B. Chatterjee, "Changing Teacher Behavior through feedback from Pupils: An Application of Equilibrium Theory," in W. W. Carters and N. L. Gage (eds.), Readings in the Social Psychology of Education, Allyn and Bacon, Boston, 1963.

8. R. C. Bryan, *Reactions to Teachers by Students, Parents and Administrators*, United States Office of Education, Cooperative Research Project No. 668, Western Michigan University, Kalamazoo, 1963.

9. M. T. Miller, "Instructor Attitudes Toward, and Their Use of, Student Ratings of Teachers," *Journal of Educational Psychology*, vol. 62, pp. 235-239, 1971.

10. Centra, "Two Studies," op. cit., p. 30.

11. L. Festinger, *A theory of Cognitive Dissonance*, Row, Peterson, Evanston, Ill., 1957.

12. Centra, "Two Studies," op. cit., p. 30.

13. Peter W. Frey, "The Ongoing Debate: Student Evaluation of Teaching," *Change*, vol. 6, no. 1, pp. 47-48, 64, 1974.

14. Ibid., p. 48.

15. F. Costin, W. Greenough, and R. Menges, "Student Ratings of College Teaching: Reliability, Validity, and Usefulness," *Review of Educational Research*, vol. 41, no. 5, pp. 511-535, 1971; and Centra, "The Student Instructional Report," op. cit.

16. Costin, Greenough, and Menges, op. cit., p. 511.

17. Ibid.

18. R. E. Spencer and L. M. Aleamoni, "A Student Course Evaluation Questionnaire," *Journal of Educational Measurement*, vol. 7, pp. 209-210, 1970.

19. F. Costin, "A Graduate Course in the Teaching of Psychology: Description and Evaluation," *Journal of Teacher Education*, vol. 19, pp. 425-432, 1968.

20. E. R. Guthrie, *The Evaluation of Teaching: A Progress Report*, University of Washington, Seattle, 1954.

21. G. M. French, "College Students' Concept of Effective Teaching Determined by an Analysis of Teaching Ratings," Dissertation Abstracts, vol. 17, pp. 1380-1381, 1957.

22. P. L. Crawford and H. L. Bradshaw, "Perception of Characteristics of Effective University Teachers. A Scaling Analysis," *Educational and Psychological Measurement*, vol. 28, pp. 1079-1085, 1968.

23. D. Musella and R. Rusch, "Student Opinion on College Teaching," *Improving College and University Teaching*, vol. 16, pp. 137-140, 1968.

24. Costin, Greenough, and Menges, op. cit., p. 515.

25. J. D. Royce, "Popularity and the Teacher," *Education*, vol. 77, pp. 233-237, 1956.

26. C. H. Weaver, "Instructor Ratings by College Students," *Journal of Educational Psychology*, vol. 51, pp. 21-25, 1960.

27. Guthrie, op. cit.

28. H. H. Remmers, *The Purdue Rating Scale for Instruction* (Manual), University Book Store, West Lafayette, Indiana, 1960.

29. Guthrie, op. cit.

30. J. A. Centra, "Percentile Norms and Mean Responses for the Michigan State University Student Instructional Rating Report," unpublished report, Office of Institutional Research, Michigan State University, 1966.

31. H. H. Remmers and D. Elliott, *The Purdue Rating Scale for Instruction* (Manual; revised edition), University Book Store, West Lafayette, Indiana, 1948.

32. R. E. Spencer, "The Course Evaluation Questionnaire: Manual of Interpretation," Office of Institutional Research, University of Illinois, 1965.

33. Ibid.; J. A. Centra, "A Michigan State University Student Instructional Rating Scale Analysis," unpublished report, Office of Instructional research, Michigan State University, 1964; and Weaver, op. cit.

34. E. W. Elliott, "Characteristics and Relationships of Various Criteria of College and University Teaching," *Purdue University Studies of Higher Education*, vol. 70, pp. 5-61, 1950; and Guthrie, op. cit.

35. Spencer, op. cit.

36. A. J. Drucker and H. H. Remmers, "Do Alumni and Students Differ in Their Attitude Toward Instructors? *Purdue University Studies of Higher Education*, vol. 70, pp. 60-74, 1950.

37. W. R. Borg, "Personality and Interest Measures as Related to Criteria of Instructor Effectiveness," *Journal of Educational Research*, vol. 50, pp. 701-709, 1957.

38. A. W. Bendig, "Ability and Personality Characteristics of Introductory Psychology Instructors Rated Competent and Empathic by Their Students," *Journal of Educational Research*, vol. 48, pp. 705-709, 1955.

39. R. L. Isaacson, W. J. McKeachie, and J. E. Milholland, "Corrections of Teacher Personality Variables and Student Ratings", *Journal of Educational Psychology*, vol. 54, pp. 110-117, 1963.

40. D. S. Cosgrove, "Diagnostic Rating of Teacher Performance," *Journal of Educational Psychology*, vol. 50, pp. 200-204, 1959.

41. J. W. Riley, A. F. Ryan, and M. Lifshitz, *The Student Looks at His Teacher*, Rutgers University Press, New Brunswick, N. J., 1950.

42. P. Seldin and E. Wakin, "How Deans Evaluate Teachers," *Change*, vol. 6, no. 9, pp. 48-49, 1974.

43. Ibid., p. 48.

44. Kerlinger, op. cit., p. 353.

45. Rodney T. Harnett, *Accountability in Higher Education: A Consideration of Some of the Problems of Assessing College Impacts*, *Educational Testing Service*, Princeton, N. J., 1971.

46. Ibid., p. 6.

47. Barry R. McGhan, "Accountability as a Negative Reinforcer," *American Teacher*, vol. 55, no. 3, November 1970, p. 13.

48. Kerlinger, op. cit. p. 353.

49. Ibid., p. 354.

50. Ibid.

51. Frey, op. cit., p. 48.

52. Ibid.

53. Arthur W. Combs and D. Snygg, *Individual Behavior* (rev. ed.), Harper & Row, New York, 1959, p. 20.

54. Seldin and Wakin, op. cit., p. 49.

55. Ibid.

Fall 1976
Volume 1 Number 1
PEPPERDINE UNIVERSITY
SCHOOL OF EDUCATION

THE commentator